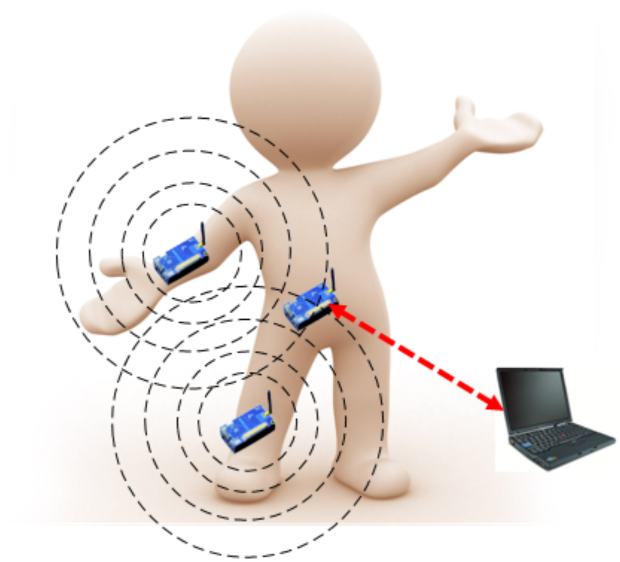
Body Area Sensor Networks II

CSE 40437/60437-Spring 2015

Prof. Dong Wang

Paper

- Paper 3: Qi, Xin, et al. "RadioSense: Exploiting Wireless Communication Patterns for Body Sensor Network Activity Recognition." RTSS. 2012.



Background - Activity Recognition

- Activity Recognition aims to automatically recognize user actions from the patterns (or information) observed on user actions with the aid of computational devices.

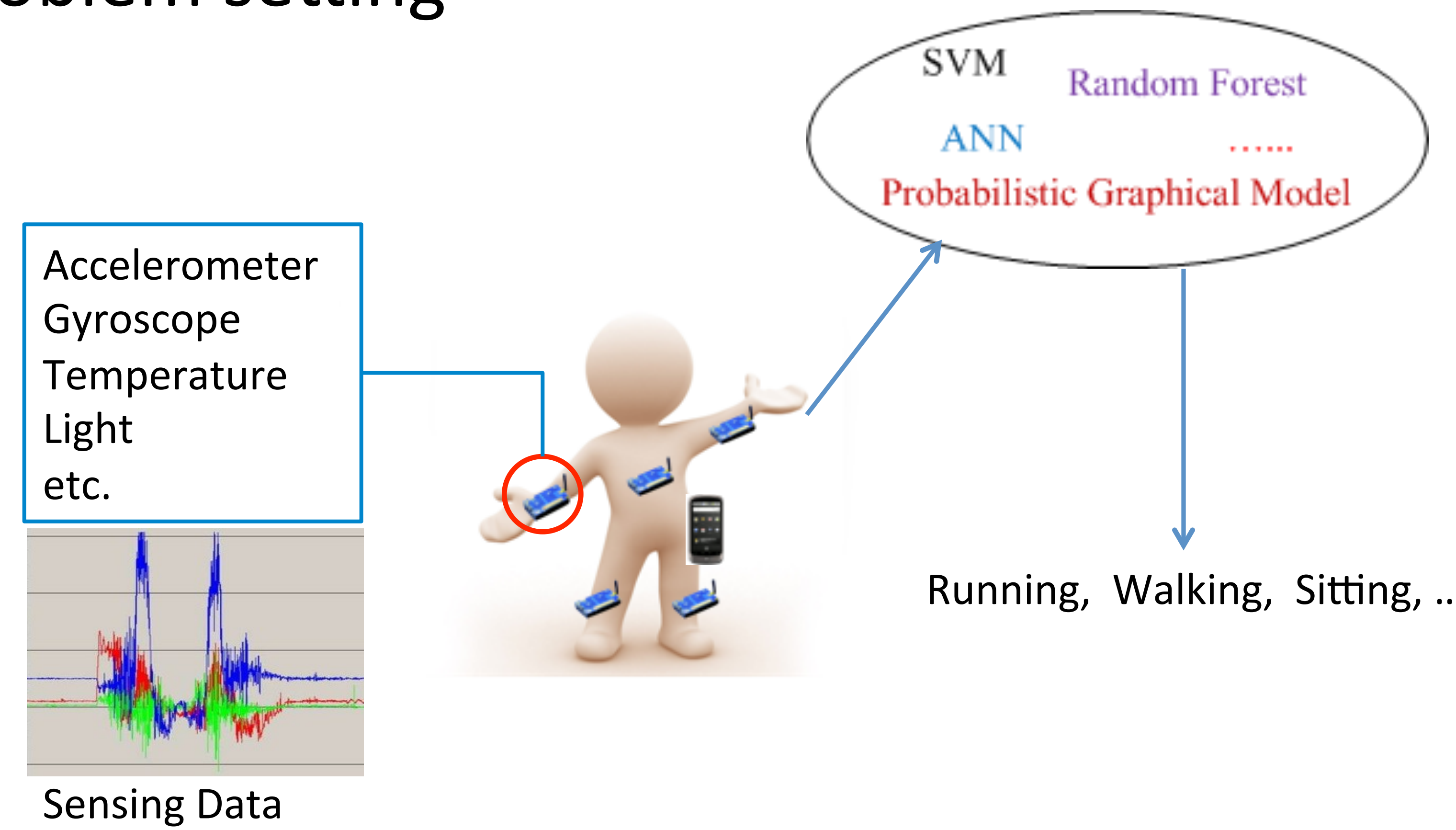
Fall Detection

Sleeping Assessment

Depression Detection

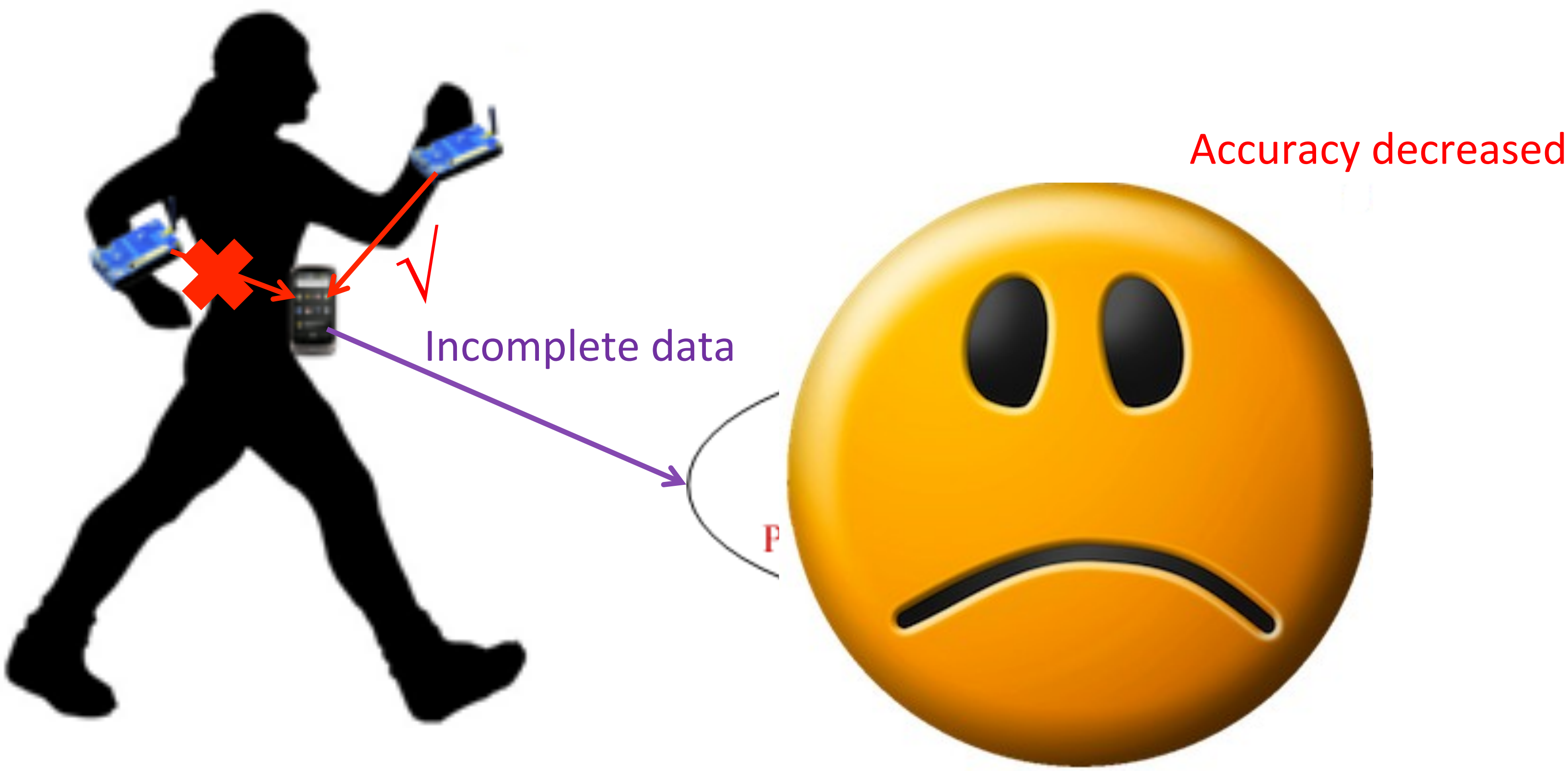
Sensing-based Activity Recognition

- Problem setting



A Dilemma – On One Hand

- Sensing data transmission suffers body fading



A Dilemma – On the Other Hand

- To increase data availability
  - Increase transmission power

Consequences:

Increase energy overheads

Increase privacy risks

Transmission Range

A Dilemma – On the Other Hand

- To increase data availability
  - Using complicated MAC protocols

Consequence:

Increase energy overheads
for retransmissions

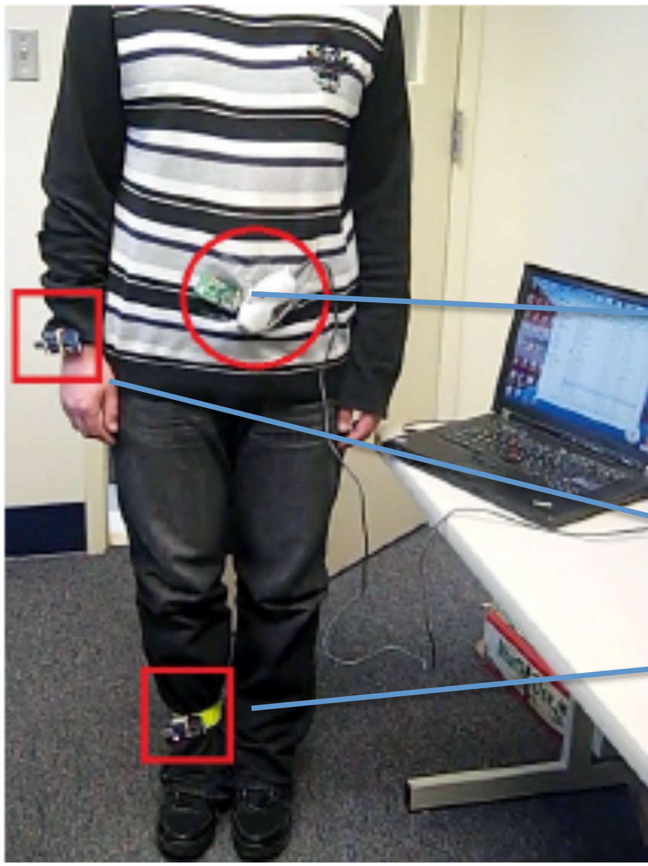
Many existing works propose new MAC protocols to improve packet delivery performance in body sensor network.

However, the impermeability of human body is a large obstacle for transmission efficiency.

The Idea and Research Question

- Idea
  - As it is difficult to overcome the impermeability of human body, can we utilize it? If so, how?
  - It is ***reasonable to imagine diff. activities have diff. patterns of packet loss and fading, which we call communication patterns.***
  - ***We use communication pattern for recognizing activities.***

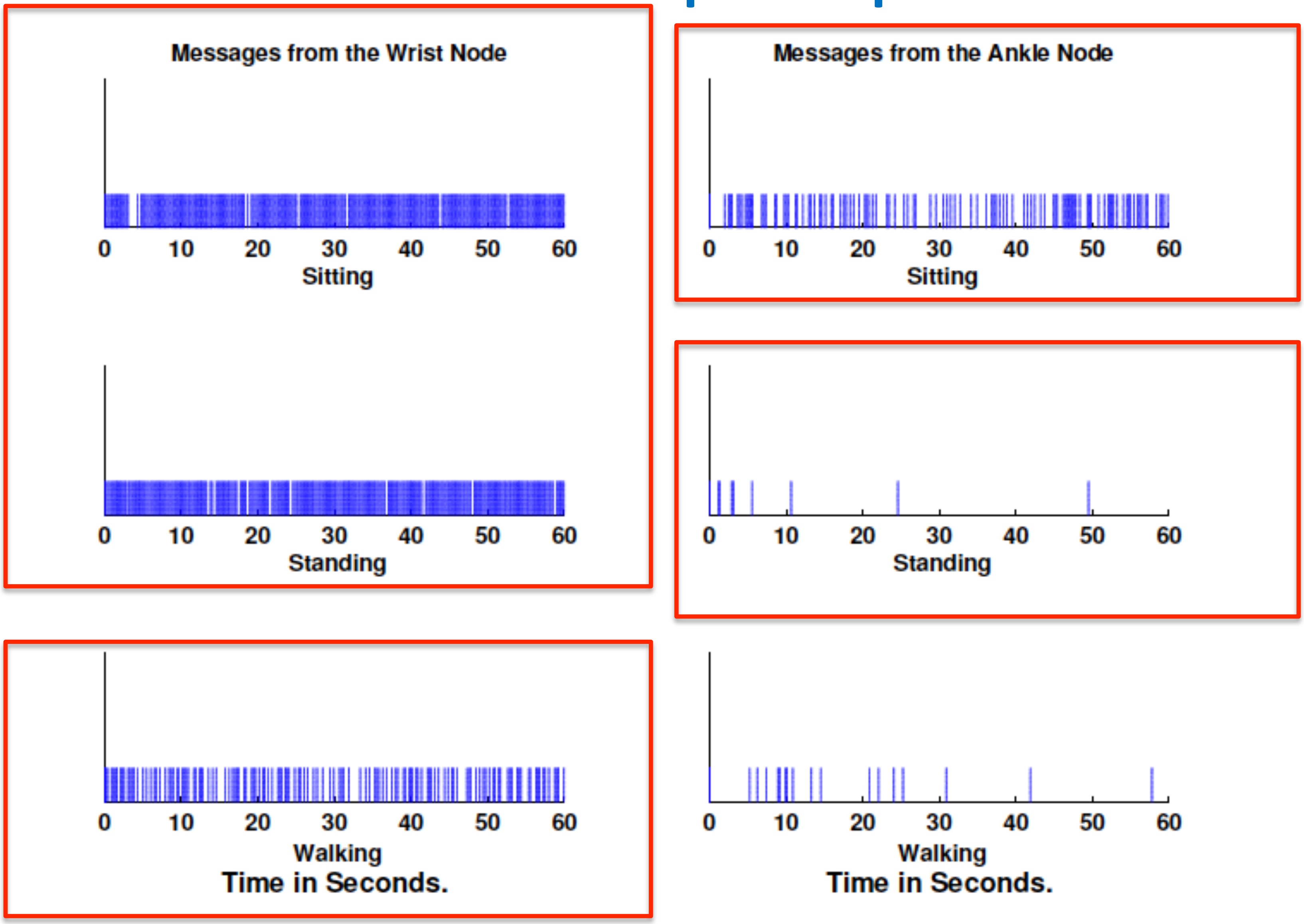
Proof of Concept Experiment



Base Station

RadioSense Nodes

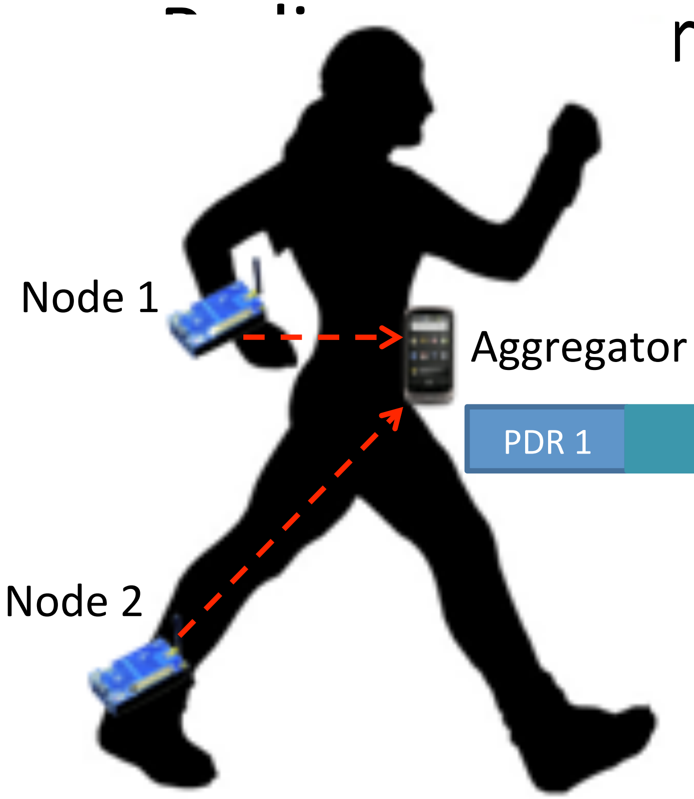
Proof of Concept Experiment

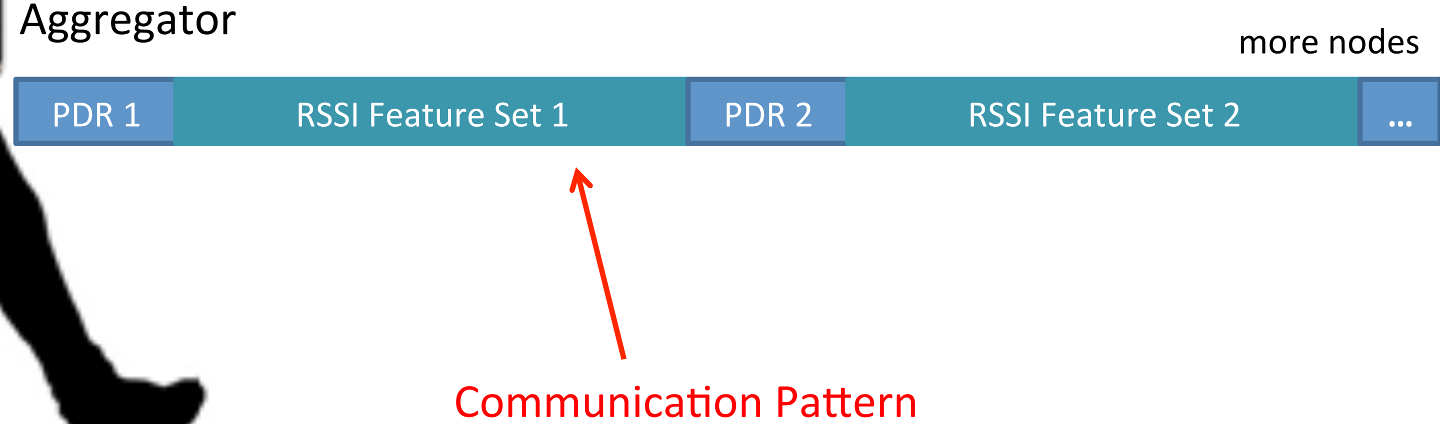


Communication Pattern

PDR = Packet Delivery Ratio

RSSI = Received Signal Strength Indicator





Factors Influencing the Discriminative Capacity of Communication Patterns

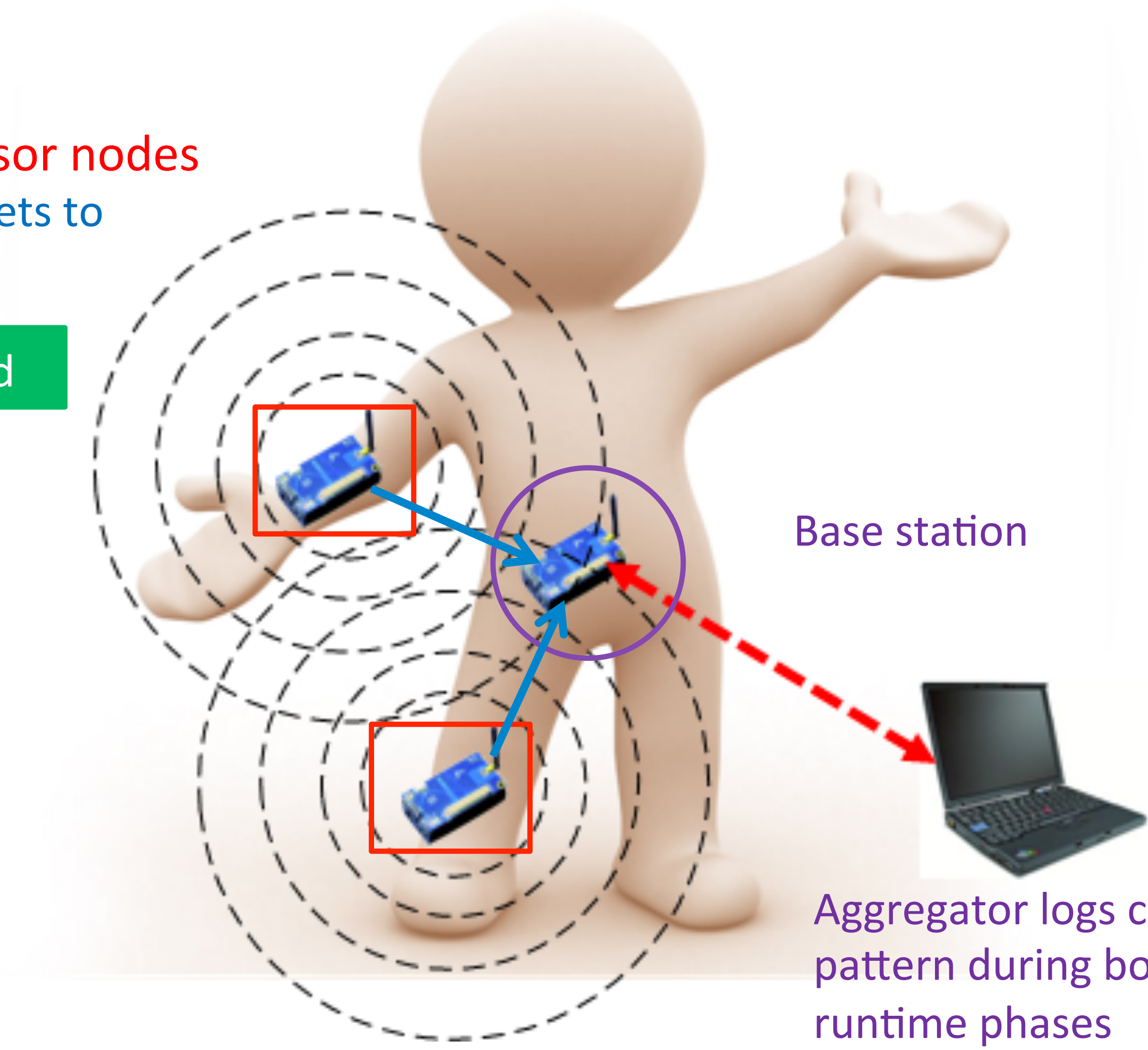
- Communication patterns
  - PDR
    - Influencing factor: transmitting power
  - RSSI features
    - Influencing factors: transmitting power, packet sending rate
  - A common influencing factor:
    - smoothing window size – length of time window for extracting features
- How to optimize the above system parameters:
  - Through benchmarking

RadioSense – a Prototype System

Two on-body sensor nodes

Send simple packets to base station

Packet::NodeId



Base station

Aggregator logs communication pattern during both training and runtime phases

Data Collection

- Aim to find insightful relationship between recognition accuracy and system parameters – one subject's data
  - Mixing multi-subjects' data may blur the relationship
- 7 activities: running, sitting, standing, walking, lying down, cycling and cleaning
  - 4-activity set: running, sitting, standing, walking
  - 6-activity set: 4-activity set + lying down and cycling
  - 7-activity set: 6-activity set + cleaning
- Transmission (TX) power level: 1~5 (maximum: 31)
- Packet sending rate: 1-4 pkts/s
- Each activity is performed for 30 minutes in diff. places (lab, classroom, living room, gym, kitchen, and outdoor)

N-fold Cross Validation

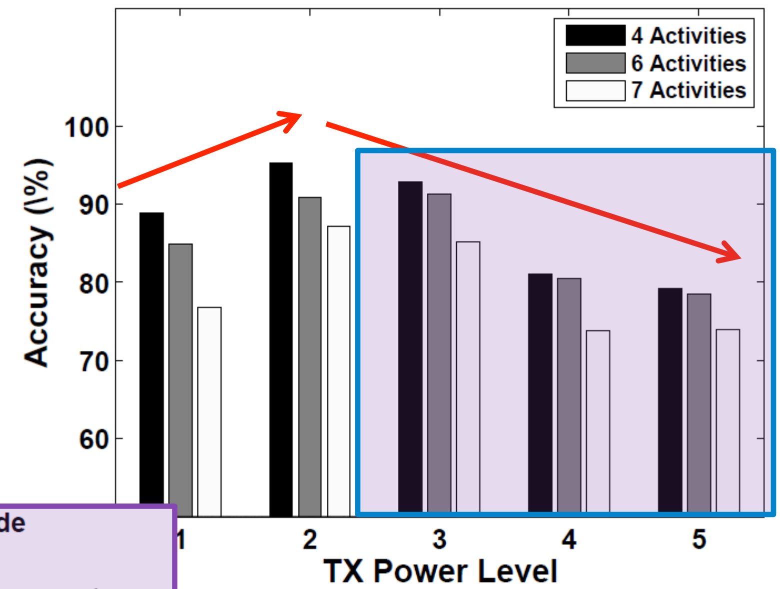
- Divide the datasets into N subsets
- N-1 subsets are used for training
- 1 subset is used for testing
- Repeat the above process for N times so that each of the N subsets is used exactly once for the testing data

TX Power Level

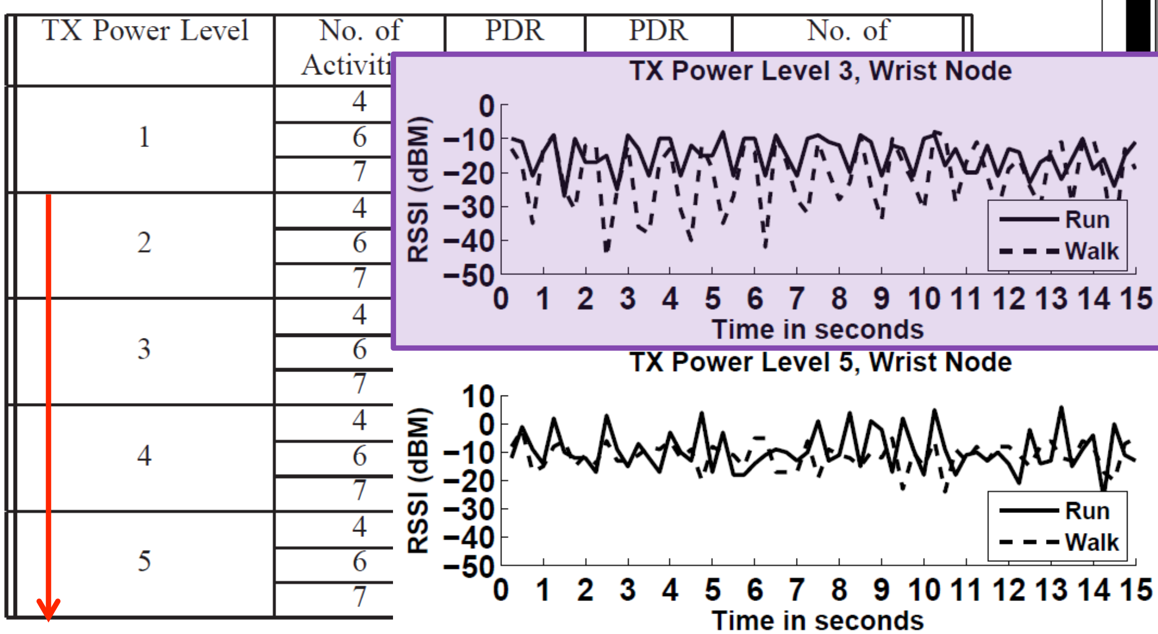
- Accuracy

first increases, then decreases

Q: Why is that?



Packet Sending Rate = 4pkts/s,
Smooth Window= 9 seconds



TX Power Level	No. of Activiti	PDR	PDR	No. of
1	4			
	6			
	7			
2	4			
	6			
	7			
3	4			
	6			
	7			
4	4			
	6			
	7			
5	4			
	6			
	7			

ower ↑

's discriminative capacity ↓

features

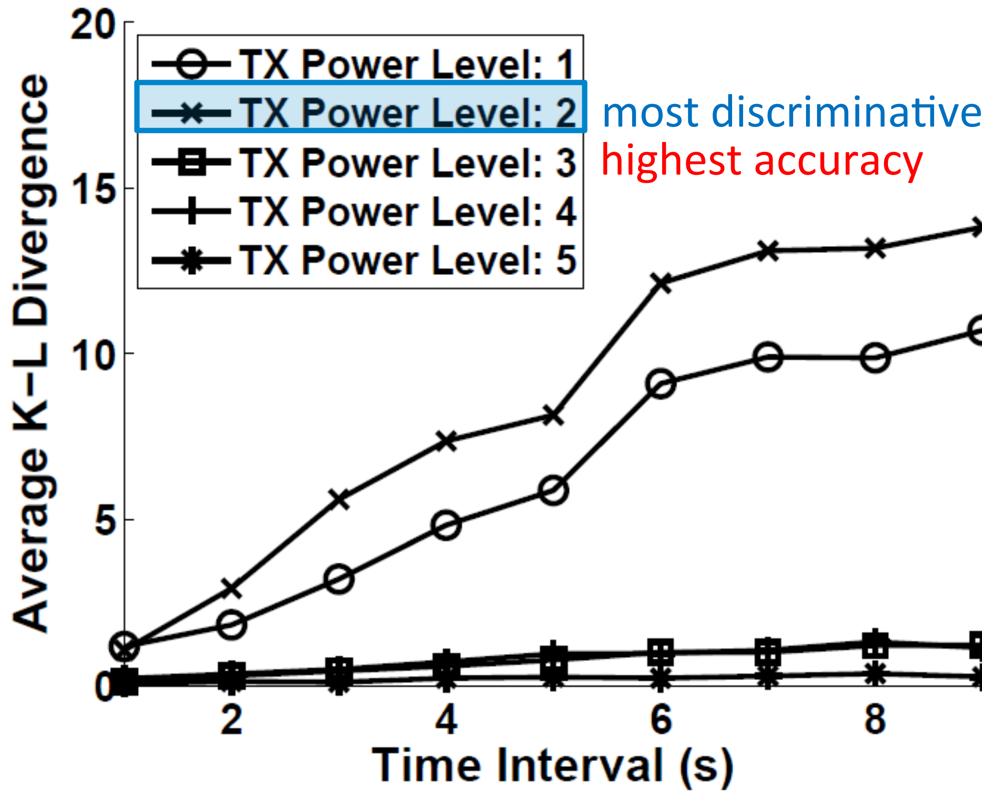
iminative capacity ↓

TX Power Level

- Quantify PDR's discriminative capacity

Metric – Average Kullback–Leibler Divergence (KLD) over all activity pairs

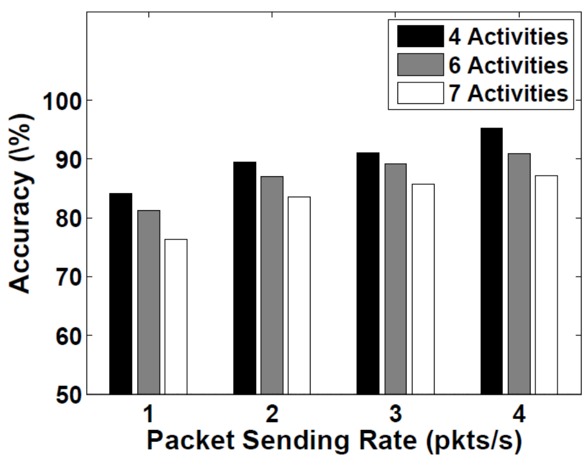
KLD: small value = similar; large value = different



Q: What do you observe from this graph?

Packet Sending Rate & Smoothing Window Size

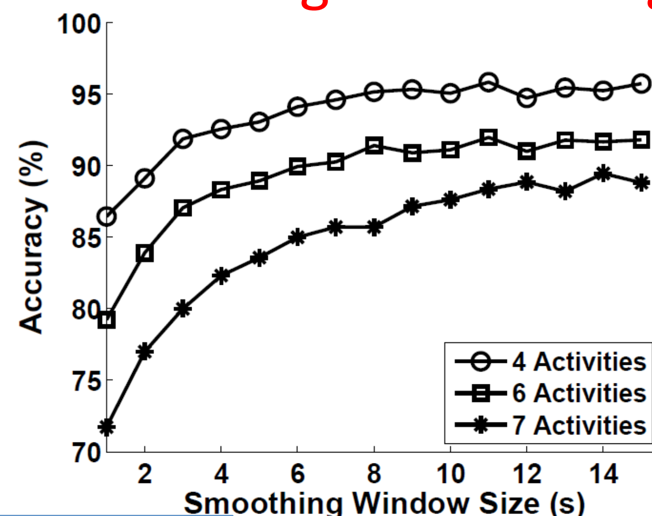
Accuracy increases with higher packet sending rate



TX Power Level = 2, Smooth Window= 9 seconds

Higher packet sending rate captures more information for RSSI variations (packet sending -> sensing the BSN channel)

Accuracy increases with larger smoothing window size



Features extracted from larger smoothing window are more robust to noise

…ending rate = 4 pkts/s

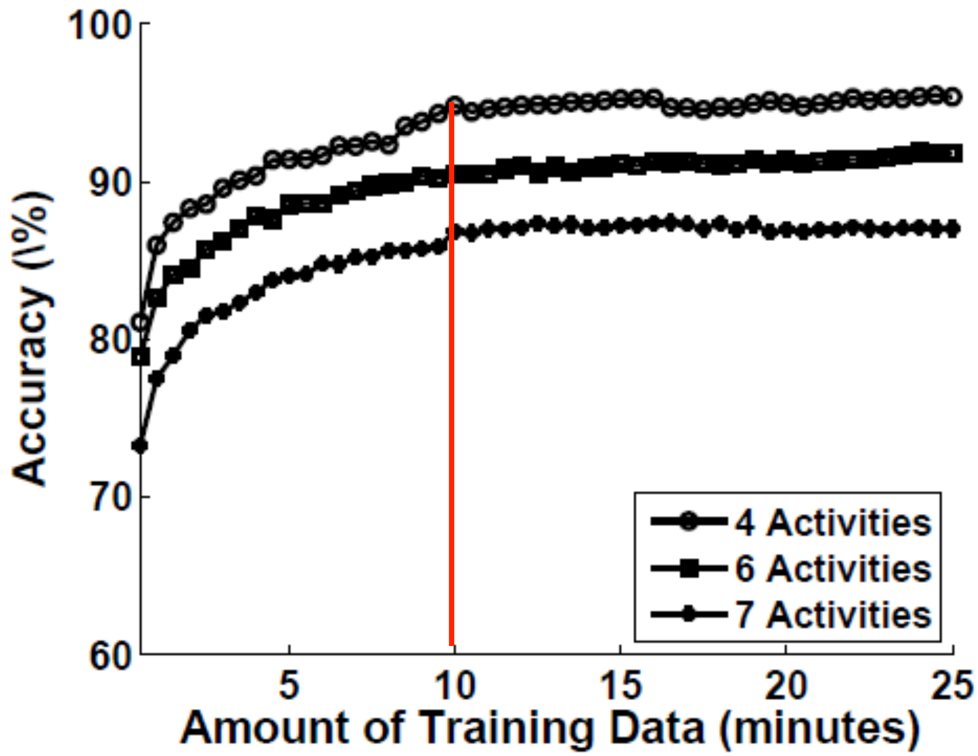
Any tradeoff you observe?

Optimize Packet Sending Rate & Smoothing Window Size

- Packet sending rate balances energy overhead and accuracy
- Smoothing window size balances latency and accuracy
- Rules for packet sending rate optimization:
  - At optimal TX power level, from 1 pkts/s, RadioSense selects i pkts/s if:
    - i achieves **90% accuracy** OR
    - **i>4**, accuracy improvement of i+1**<2%**
- Rules for smoothing window size optimization:
  - At optimal TX power level and packet sending rate, RadioSense selects i seconds if:
    - i achieves **90% accuracy** OR
    - **i>10** seconds, accuracy improvement of i+1**<2%**

Amount of Training Data

- Average accuracy of three subject with different amount of training data
- 10-minute data is enough for stable accuracy

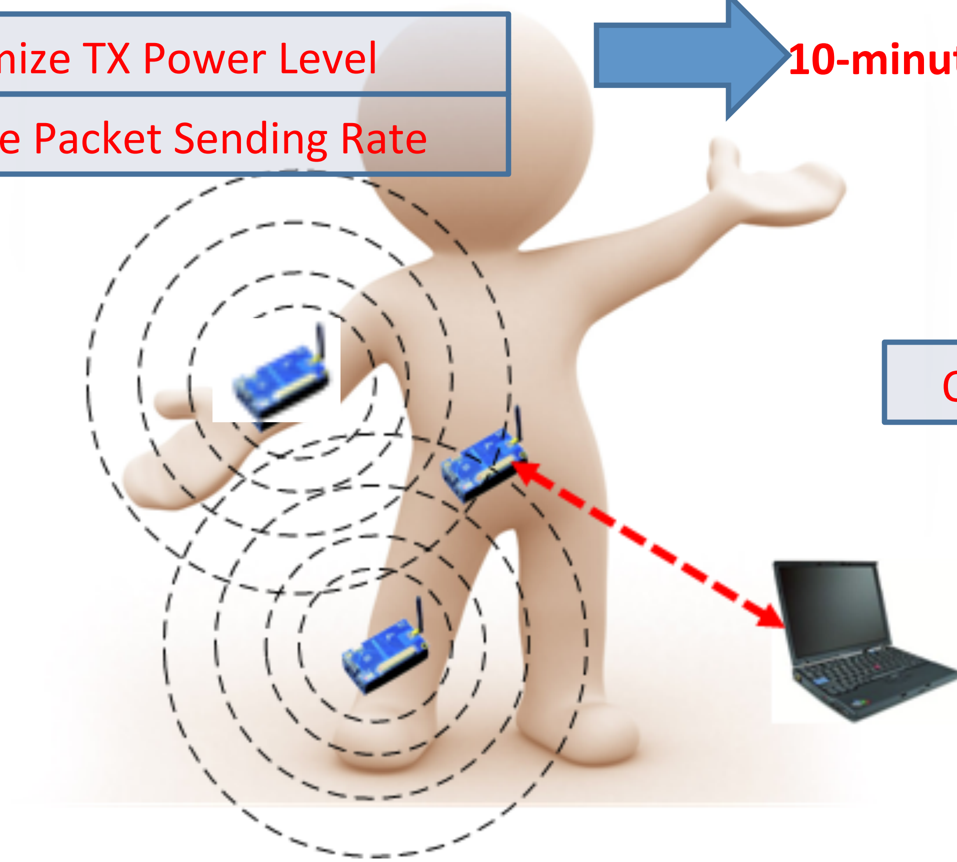


RadioSense Recap

- In training phase, we design RadioSense to bootstrap the system following the steps below:

Optimize TX Power Level

Optimize Packet Sending Rate



10-minute training data for each activity

Optimize Smoothing Window Size

Trained Classifier

Up to Now

- We have answered …
  - How to endow the communication pattern with enough discriminative capacity for recognizing diff. activities?
- In the evaluation, we will answer…
  - What are the impacts of using communication pattern for AR on other system performance issues, such as energy and privacy?

Evaluation – Data Collection

- 3 subjects

Subject	Gender	Height (m)	Weigh (kg)
1	Male	1.80	85.0
2	Male	1.68	63.0
3	Female	1.56	48.0

- 7 activities - running, sitting, standing, walking, lying, riding and cleaning
- Different places - lab, classroom, living room, gym, kitchen, and outdoor
- During training phase
  - Each subject performs activities for system parameter optimization
  - With the optimal parameters, for each activity, each subject collects 10-minute data for training and 30-minute data for testing
- Lasts for two weeks
- One classifier for each subject

Evaluation – System Parameter Optimization

- Average KLD Table

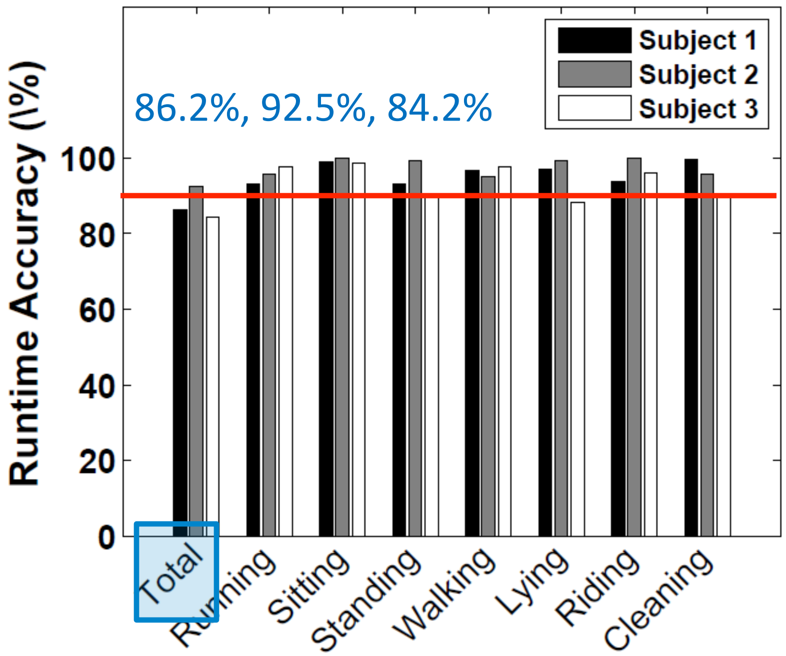
	Subject 1	Subject 2	Subject 3
TXPowerLevel 1	9.88	12.86	**12.90**
TXPowerLevel 2	**13.18**	**20.79**	8.80
TXPowerLevel 3	1.21	1.59	1.58

Subject 3 is smaller than the other two

Subject	Gender	Height (m)	Weigh (kg)
1	Male	1.80	85.0
2	Male	1.68	63.0
3	Female	1.56	48.0

Subject	TX Power Level	Packet Sending Rate (pkts/s)	Smoothing Window Size (seconds)
1	2	4	8
2	2	3	6
3	1	4	6

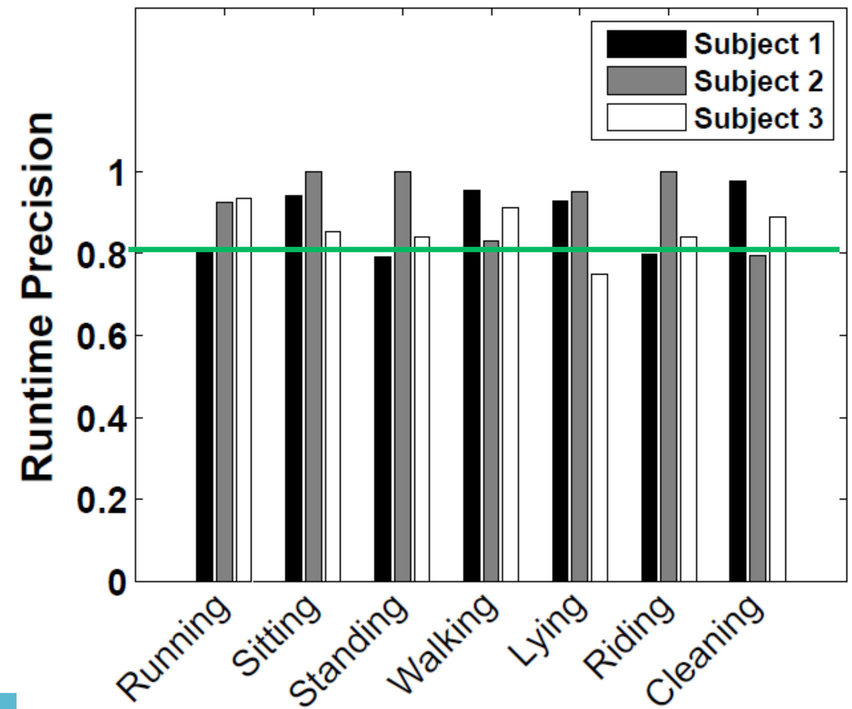
Evaluation – Accuracy and Precision



Most single activity achieves 90% accuracy

Ave. KLD: 13.18, 20.79, 12.90

Ave. KLD is a validated metric



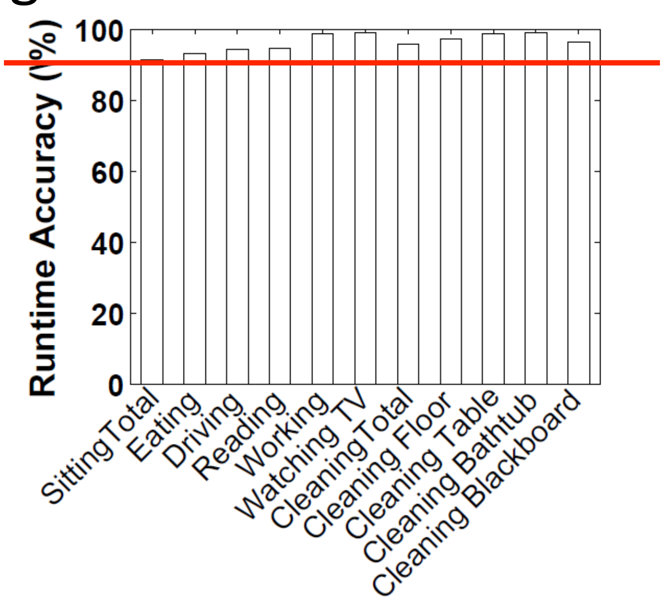
Most single activity achieves 0.8 precision

Accuracy = (TP+TN)/(TP+TN+FP+FN)

Precision = TP/(TP+FP)

Potentials – More Fine-Grained Activities

- One subject
- **Sitting set** - driving, working, reading, eating, and watching TV
- **Cleaning set** - cleaning table, cleaning floor, cleaning bathtub, and cleaning blackboard
- 10-minute data of each activity for training, 30-minute data of each activity for testing
- **Sitting - 91.5%**
- **Cleaning - 95.8%**



Evaluation – Battery Lifetime and Privacy

- Battery lifetime – for each subject's optimal system parameters
  - 3 Tmotes with new batteries (AA, Alkaline, LR6, 1.5V)
  - Run RadioSense until batteries die
  - **159.3 hours, 168.7 hours, 175.3 hours**

Packet::NodeId

- Privacy

Lower TX power and smaller communication range: around 1 m

TX Power Level	TX Power (dBm)	Max Comm. Diameter (cm)
1	-33.0	77.2
2	-28.7	108.3
3	-25.0	388.0
7	-15.0	923.2

Commonly used TX power level in BSN, [Mobicom '09]

Privacy risks are reduced!

Evaluation - Potential of Coexistence with Other On-body Sensor Nodes

- RadioSense - **two dedicated on-body sensor nodes**, right wrist and ankle, with optimal parameters
- **Two general purpose on-body sensor nodes**, left wrist and ankle, TX power level 7, packet sending rate 4 pkts/s
- One subject, for each activity, 10-minute training data, 30-minute testing data
- For general purpose nodes:
  - **PDR: 98.0%, 95.6%**
  - In good condition [Sensys '08], since interference from RadioSense nodes is low
- For RadioSense nodes:
  - **Accuracy: 90.8% (with other nodes), 86.3% (without other nodes)**
  - **Communication contention with other on-body nodes may amplify the discriminative capacity of communication pattern**

RadioSense does not affect general purpose nodes

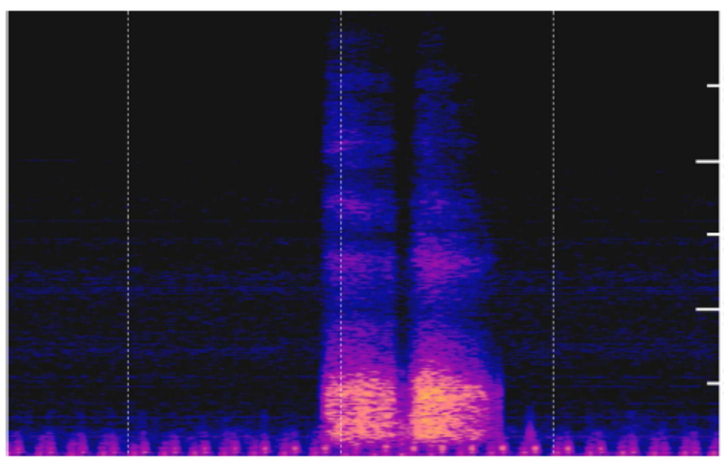
RadioSense leverages interference rather than suffers from it!

Limitations

- Strong background noises
  - Sensing-based approaches will also fail in such case because of high packet loss.
- Not scalable for new activities
  - It is a common problem for AR system using supervised learning method
- Current system is a little bit clunky
  - In future, the authors may replace the aggregator with smartphone;

Paper

- Paper 4: Yatani, Koji, and Khai N. Truong. "Bodyscope: a wearable acoustic sensor for activity recognition." Proceedings of the 2012 ACM Conference on Ubiquitous Computing. ACM, 2012.



User Activity Recognition

GPS

Accelerometers

RFID

Camera

Microphone

Bluetooth

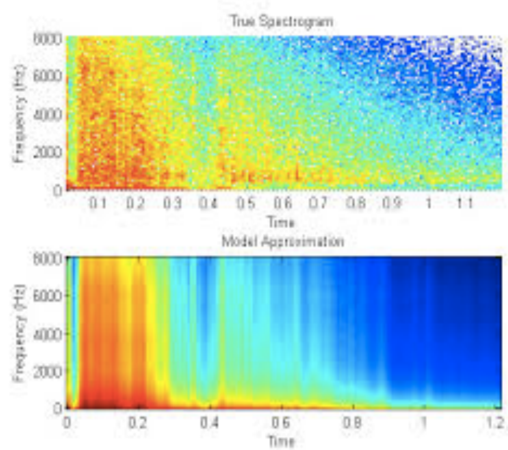
Cellular

WiFi

Q: Can we use a **single type** of sensor to detect a rich set of user's activity?

Main Idea

- Sound produced in user's throat area -> User Activities (e.g., drinking, eating, laughing, speaking, etc.)



Drinking

Eating

Speaking

Laughing

.
.
.

Coughing

Share Your Thoughts

- Q1: How would you design a system to leverage user's sound to detect different activities (e.g., drinking, eating, speaking, laughing, coughing, sighing, etc.)?

- Q2: What are the equipment(s) you might need to implement your system?

BodyScope

Bluetooth headset

Uni-directional Microphone

Chestpiece of a stethoscope

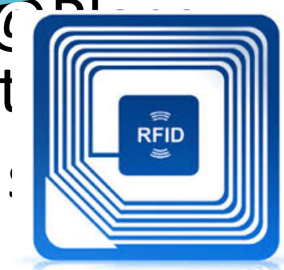
BodyScope

12 Sound-related User Activities

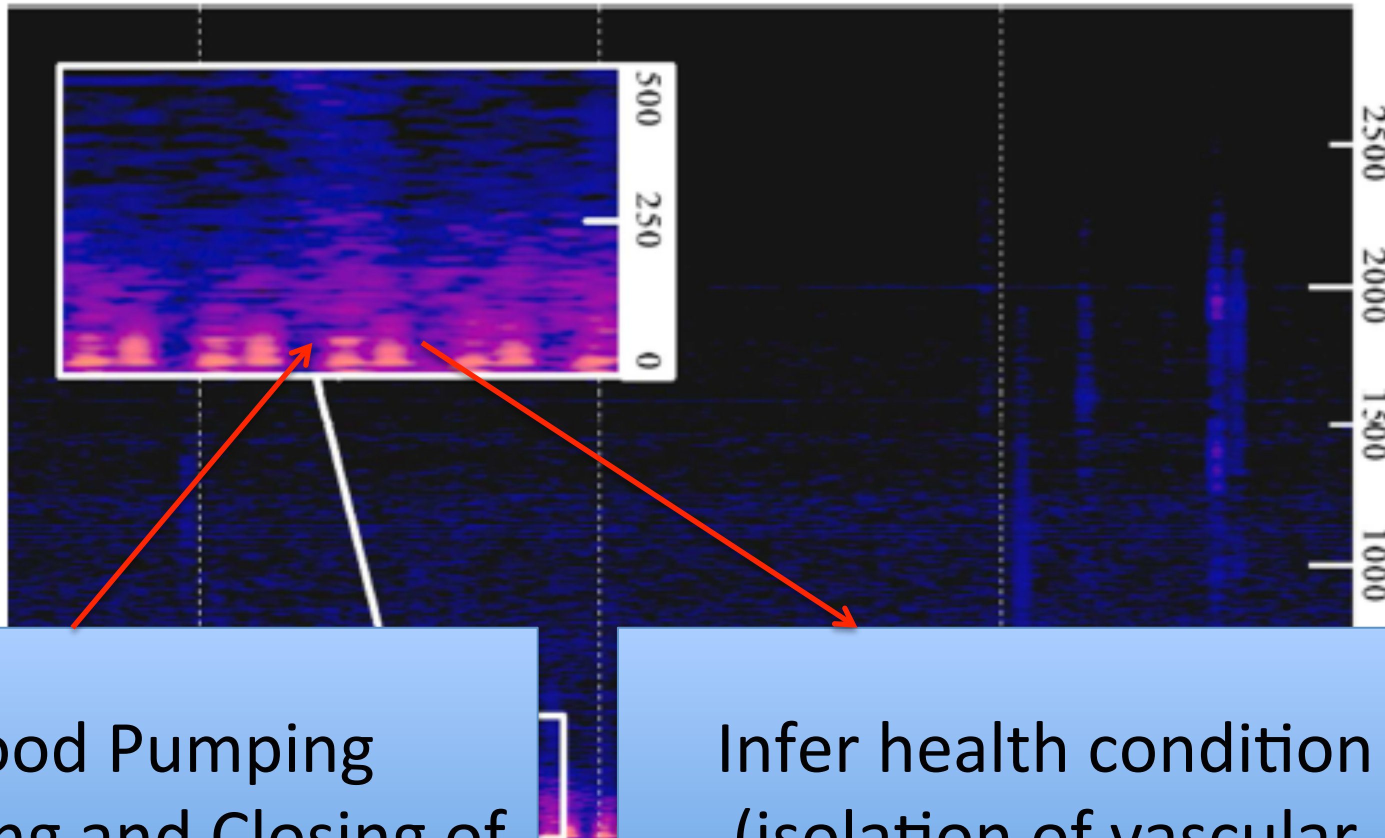
State-of-the-arts

- **GPS sensors:** infer activities related to locations (e.g., working, shopping, driving, etc.)
- **Accelerometers:** recognize user's p… movement activities (… runni…
- …sense@… …places and infer user act…
- **…ombination of the above+ WiFi:** ACE to … context of a user and infer the activities
- **RFID (Radio Frequency Identification)**: Embed RFID readers to smart gloves and install RFID tags to objects. Infer user activities from interactions (e.g., washing hands, preparing food or a drink, etc.)

Use Microphone Sensor Alone to Detect a Rich Set of User Activities!



User Activities Detected Through Sound
Sound Spectrogram



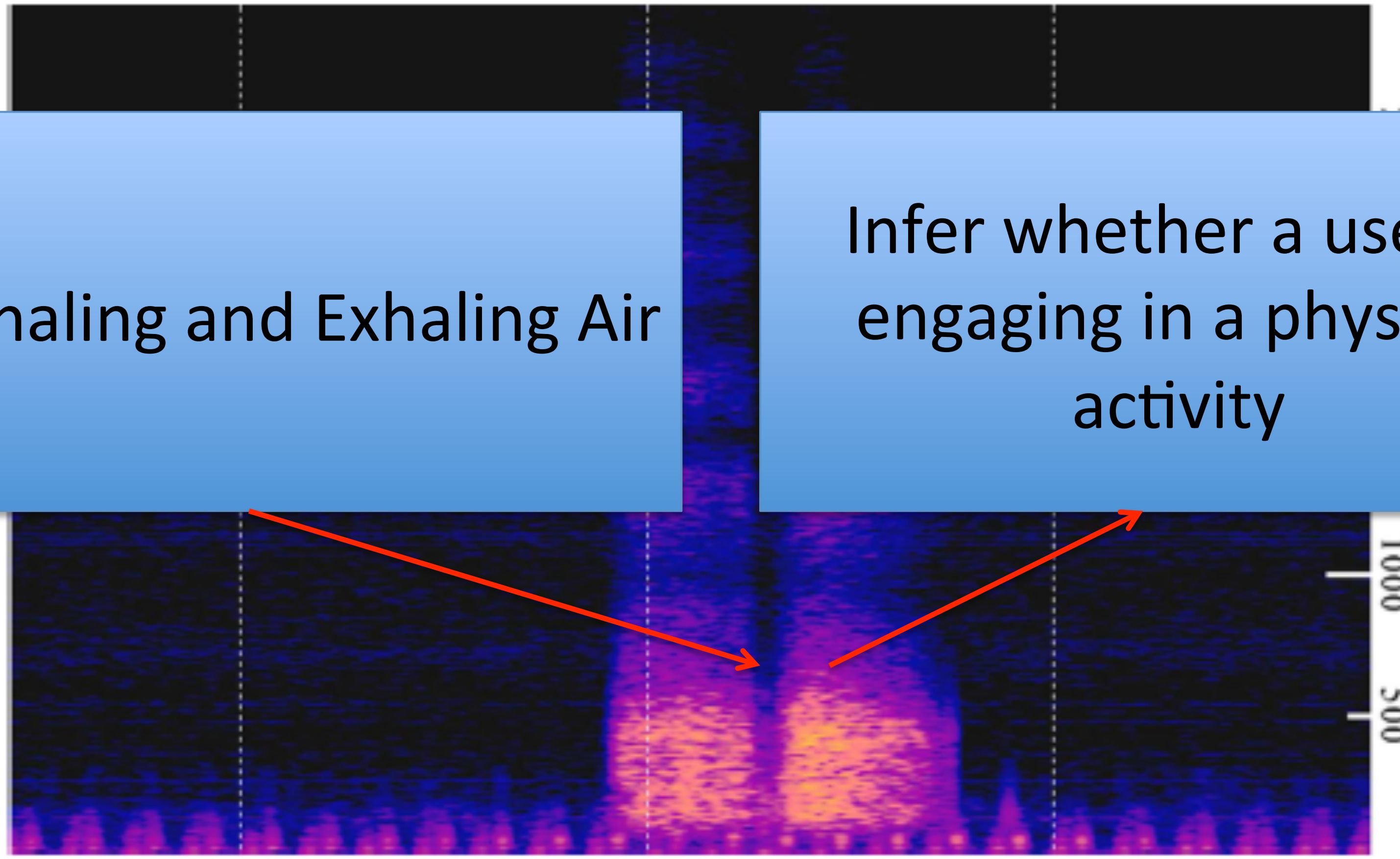
Blood Pumping
(Opening and Closing of
Atrioventricular Valves)

Infer health condition
(isolation of vascular
sound is challenging)

User Activities Detected Through Sound

Inhaling and Exhaling Air

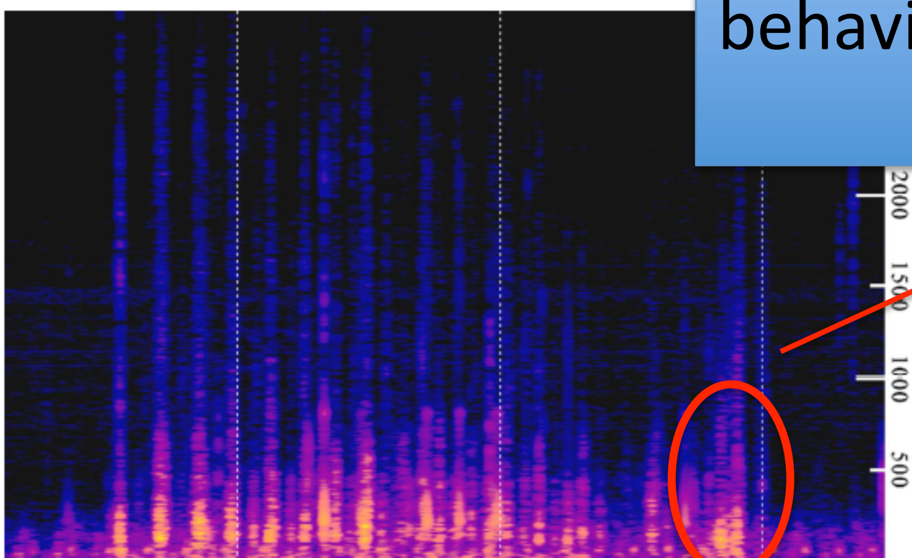
Infer whether a user is engaging in a physical activity



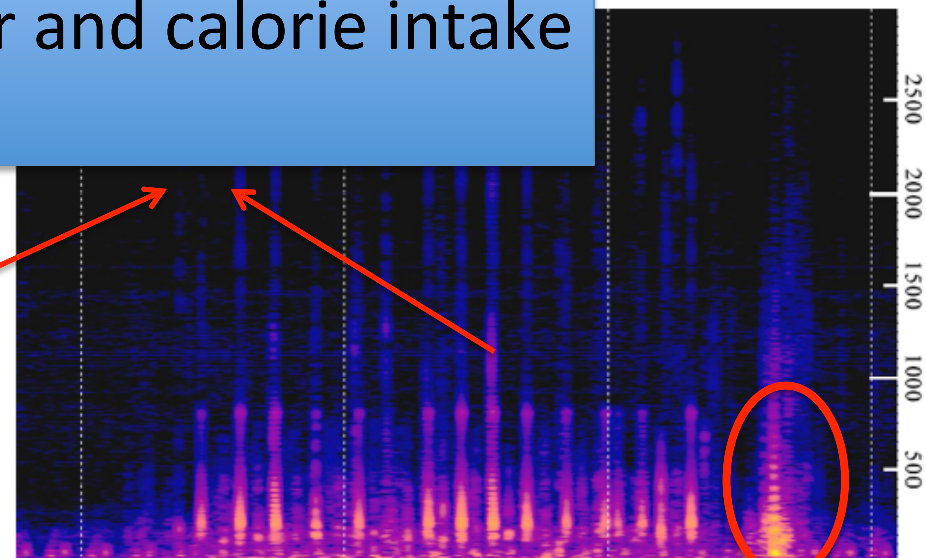
Deep Breath

User Activities Detected Through Sound

Infer user's dietary behavior and calorie intake



Cookie



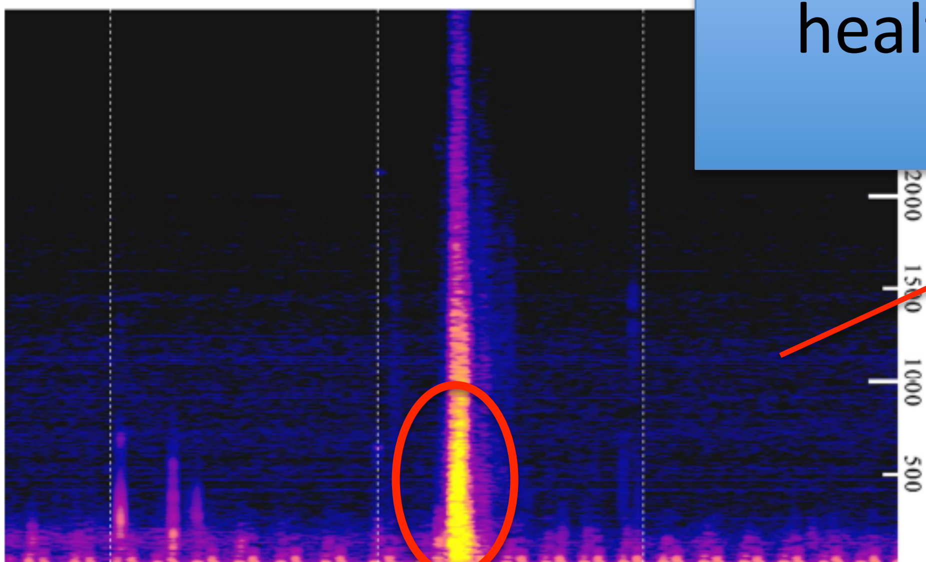
Bread

Swallowing Sound

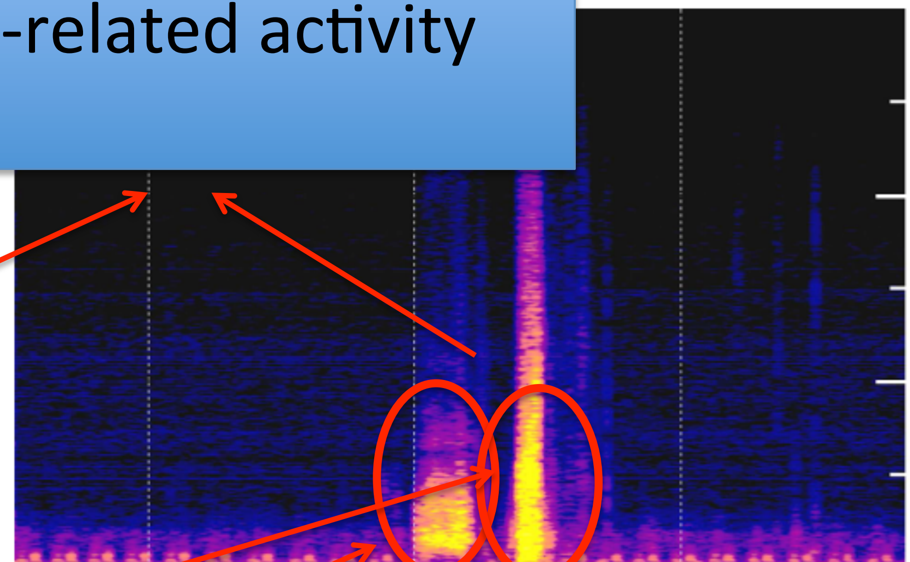
kie)

User Activities Detected Through Sound

Infer user's fluid intake and health-related activity



Cold Water



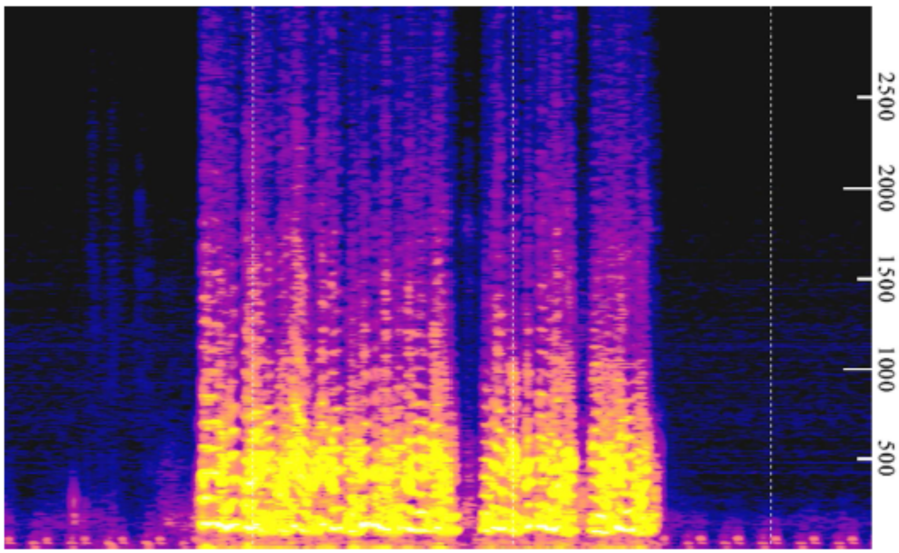
Hot Tea

Gulping

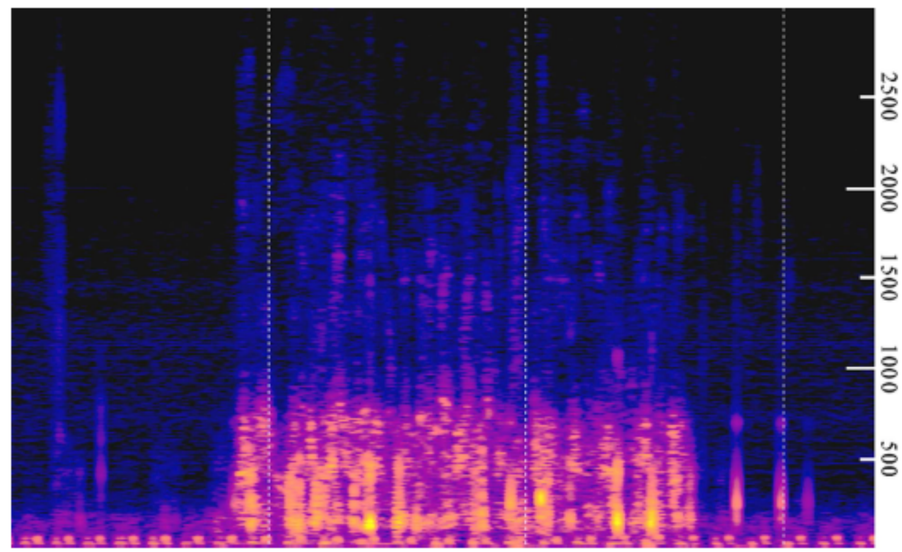
Sipping

g (Wal)

User Activities Detected Through Sound



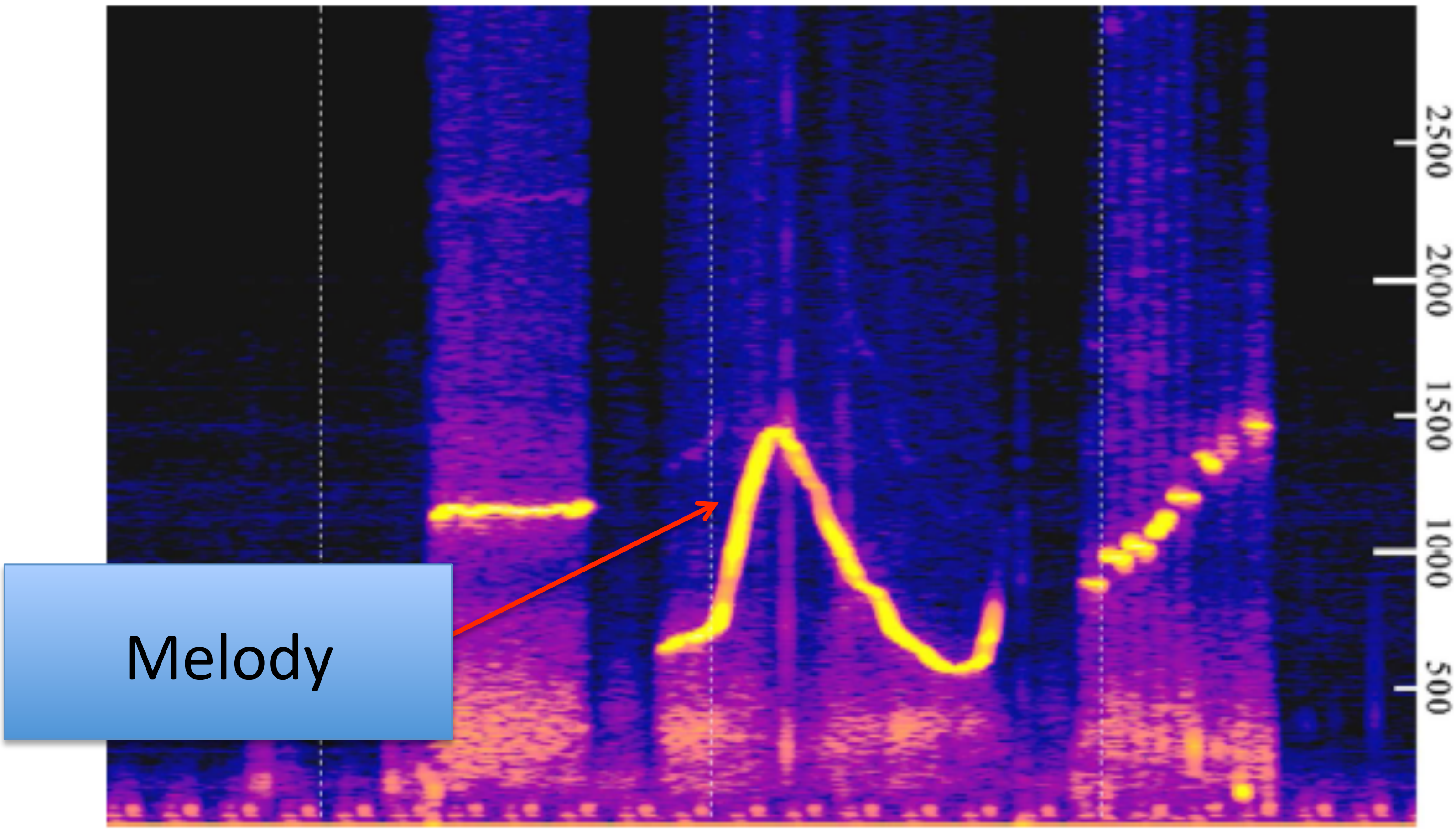
Speaking



Whispering

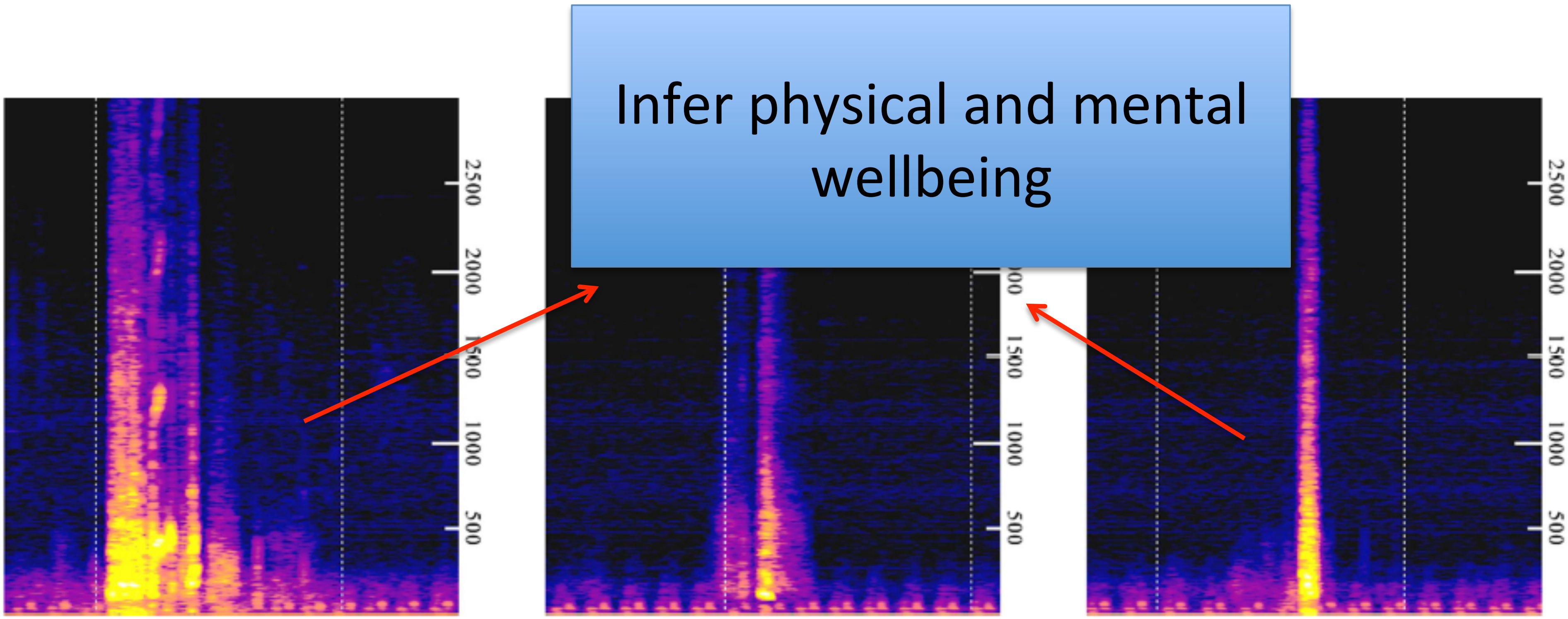
Speaking (Speaking vs Whispering)

User Activities Detected Through Sound



Whistling

User Activities Detected Through Sound



Laughing

Sighing

Coughing

Non-verbal Sounds (Laughing, Sighing, Coughing)

Classification Technique

- Features:
- Time-domain Feature
  - Zero-crossing Rate (ZCR): differentiate voiced and unvoiced sound
- Frequency-domain Feature
  - Total Spectrum Power
  - Brightness
  - Spectral Rolloff and Flux
- Classifier: SVM, Naïve Bayes, k-NN

Lab Evaluation

- Activities to be classified (12 in total):
  Seated,
  Deep breath,
  Eating cookies, Eating bread,
  Drinking, Drinking with a sip;
  Speaking; Whispering;
  Whistling;
  Laughing, Sighing, Coughing

Lab Evaluation

- Participants:
  - 10 participants (9 male, 1 female, all in 20s and 30s), all in good health
- Training and Testing Procedure
  - Leave-one-participant-out cross validation
    - Use the data from 9 participants for training and the data from the other participant for 1 test
  - Leave-one-sample-per-participant-out cross validation
    - Reserve one sample for one class from each participant as a test case and use the rest for training

Lab Evaluation

	Leave-one-participant-out			*Leave-one-sample-per-participant-out*		
	PR	**RE**	**F**	**PR**	**RE**	**F**
Bayes	47.0%	45.7%	46.3%	72.3%	71.2%	72.2%
5-NN	43.5%	43.2%	43.3%	75.3%	75.1%	75.2%
SVM	50.2%	49.1%	49.6%	79.6%	79.4%	79.5%

Training the modeling with user specific samples will be helpful!

Comparison of two validation approaches

Lab Evaluation

- Q: What activities do you think that are more difficult to be distinguished from each other?

  Seated,

  Deep breath,

  Eating cookies, Eating bread,

  Drinking, Drinking with a sip;

  Speaking; Whispering;

  Whistling;

  Laughing, Sighing, Coughing

Lab Evaluation

Actual Activities	Prediction												
	Seated	Deep breath	Eating (Cookie)	Eating (Bread)	Drinking	Drinking (with a sip)	Speaking	Whispering	Whistling	Laughing	Sighing	Coughing	**Recall [%]**
Seated	61	9	1	2	20	4	0	0	0	0	3	0	**61.0**
Deep breath	2	15	9	7	7	2	0	16	4	15	21	2	**15.0**
Eating (Cookie)	0	2	56	20	2	4	0	9	0	4	2	1	**56.0**
Eating (Bread)	2	4	27	51	5	2	0	0	1	3	1	4	**51.0**
Drinking	8	8	4	3	35	16	0	0	1	2	20	3	**35.0**
Drinking (with a sip)	3	10	17	10	33	9	0	3	0	6	6	3	**9.0**
Speaking	0	0	0	0	0	0	90	4	0	3	0	3	**90.0**
Whispering	0	11	5	0	2	1	20	53	2	3	0	3	**53.0**
Whistling	0	1	0	0	0	2	0	1	96	0	0	0	**96.0**
Laughing	1	14	4	1	6	4	8	7	1	46	3	5	**46.0**
Sighing	7	21	5	11	10	0	1	0	0	8	28	5	**28.0**
Coughing	4	3	2	3	5	3	2	0	1	11	4	62	**62.0**
Precision [%]	**69.3**	**15.3**	**43.1**	**50.5**	**27.8**	**15.8**	**75.0**	**56.4**	**90.6**	**45.5**	**31.8**	**68.1**	

The confusion matrix of the classification with **leave-one-participant-out** protocol (SVM)

Lab Evaluation

Actual Activities	Prediction												Recall [%]
	Seated	Deep breath	Eating		Drinking		Speaking	Whispering	Whistling	Laughing	Sighing	Coughing	
Seated	94	0	0	0	4	1	0	0	0	1	0	0	**94.0**
Deep breath	0	79	0	2	3	2	0	5	0	4	2	3	**79.0**
Eating (Cookie)	0	1	81	7	3	6	0	1	0	1	0	0	**81.0**
Eating (Bread)	0	1	8	80	4	5	0	0	0	0	1	1	**80.0**
Drinking	0	5	3	1	78	5	0	1	0	2	5	0	**78.0**
Drinking (with a sip)	2	2	10	5	14	60	0	2	0	1	2	2	**60.0**
Speaking	0	0	0	0	0	0	97	0	0	2	0	1	**97.0**
Whispering	0	6	2	0	0	0	4	82	0	5	0	1	**82.0**
Whistling	0	1	0	0	0	0	0	1	98	0	0	0	**98.0**
Laughing	0	7	3	0	5	4	4	6	0	64	2	5	**64.0**
Sighing	2	10	0	4	6	1	0	1	0	6	66	4	**66.0**
Coughing	4	4	1	1	1	2	0	1	0	4	8	74	**74.0**
Precision [%]	**92.2**	**68.1**	**75.0**	**80.0**	**66.1**	**69.8**	**92.4**	**82.0**	**100**	**71.1**	**76.7**	**81.3**	

Decreasing the activity granularity would help improve accuracy as well.

The confusion matrix of the classification with **leave-one-sample-per-participant-out** protocol (SVM)

Small Scale In-the-Wild Evaluation

- Participants:
  - 5 participants (3 male, 2 female,)
- Focus on 4 activities
  - Eating, drinking, speaking and laughing
- Ground-truth:
  - Ask users to wear another phone with cameras around neck to record user's activities

Small-scale In-the-Wild Evaluation

Actual Activities	Eating	Drinking	Speaking	Laughing	Recall [%]
Eating	157	11	11	0	**87.8**
Drinking	19	33	7	9	**56.0**
Speaking	16	10	498	7	**93.8**
Laughing	1	0	25	14	**35.0**
Precision [%]	**81.3**	**61.1**	**92.1**	**66.7**	

The confusion matrix of the **SVM** classification in our small-scale in-the-wild study

Small-scale In-the-Wild Evaluation

		Prediction: Eating	Prediction: Drinking	Prediction: Speaking	Prediction: Laughing	**Recall [%]**
Actual Activities	Eating	125	49	4	1	**69.8**
	Drinking	17	36	6	0	**61.0**
	Speaking	58	40	352	81	**66.3**
	Laughing	1	1	16	22	**55.0**
Precision [%]		**62.2**	**28.6**	**93.1**	**21.2**	

The confusion matrix of the **Naïve Bayes classification** in our small-scale in-the-wild study

A Demo Video

- BodyScope on Youtube:
- https://www.youtube.com/watch?v=ns-BIh8p8IU

Limitation and Future Work

- Limited accuracy of the prototype
  - F-measure: 79.5% for lab experiment and 71.5% for in-the-wild study
- Users need to wear a special device that is very visible
  - Build a more comfortable and less intrusive device
- Only 12 activities are studied
  - Sense more activities (e.g., smoking, sneezing, etc.)
- Privacy issues have not been studied
  - Voice and sound contains a lot of sensitive and personal information

Thank You!

The End.